ARL-TR-8770 ● SEP 2019

DEVCOM
ARMY RESEARCH LABORATORY

Human–Robot Interaction Design Research: From Teleoperations to Human–Agent Teaming

by Michael Barnes, Linda R Elliott, Julia Wright, Angelique Scharine, and Jessie Chen

NOTICES

Disclaimers

ARL-TR-8770 ● SEP 2019

Human–Robot Interaction Design Research: From Teleoperations to Human–Agent Teaming

by Michael Barnes, Julia Wright, Angelique Scharine, and Jessie Chen
Human Research and Engineering Directorate, CCDC Army Research Laboratory

Linda R Elliott
CCDC Data & Analysis Center

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
September 2019

2. REPORT TYPE
Technical Report

3. DATES COVERED (From - To)
October 2018–August 2019

4. TITLE AND SUBTITLE
Human–Robot Interaction Design Research: From Teleoperations to Human–Agent Teaming

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Michael Barnes, Linda R Elliott, Julia Wright, Angelique Scharine, and Jessie Chen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CCDC Army Research Laboratory
ATTN: FCDD-RLH-BD
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-8770

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID(s): Angelique Scharine, 0000-0001-8230-1824; Julia Wright, 0000-0003-3026-1538

14. ABSTRACT
This report covers human factors research supported by the Army Research Laboratory's human–robot interaction (HRI) multiyear program. The emphasis is on design principles derived from the HRI research commencing with early research on teleoperations to current research on human–agent teaming (HAT), including topics such as human factors of teleoperations, multimodal control and display research, adaptive systems, RoboLeader (a planning agent), models of trust and transparency, design of visualization for effective HAT, and a discussion of current and future efforts related to bidirectional communications between humans and agents. An important principle underlying our research is that as agents become more autonomous, Soldiers' taskings do not become easier—rather they change, often becoming more demanding.

15. SUBJECT TERMS
human–agent interactions, teleoperations, multimodal interfaces, transparency, trust, visualization, bi-directional communications

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Michael Barnes
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	UU	54	(520) 538-4702

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

1. Introduction

The military is going through a change in strategic emphasis from asymmetric warfare threats involving intact terrorist groups and state-sponsored disruptive insurgent networks to focusing on nation states. Existential future threats are perceived to come from near-peer nations (Russia and China) and unstable nation states such as North Korea and Iran (Washington Times 2019). The unsettling feature of these threats is that some of the actors are supported by military forces as large as or larger than ours and their technology is continuing to improve to the point that absolute technological advantage cannot be assumed. In particular, China's progress in robotics and artificial intelligence (AI) will impact both their commercial and military strength. AI will not only improve response time and precision of weapons but autonomous systems will act as force multipliers (Barnes and Chen 2012; Chen and Barnes 2014; Defense Science Board 2016). Modern forces will consist of manned, unmanned, and autonomous elements controlled directly or indirectly by human supervisors (Goodrich and Schultz 2007; Evans et al. 2017; Scharre 2018). For example, developing systems such as the Next Generation Combat Vehicle robotic version (NGCV [RV]) combines all three elements and types of control to maximize synergy between human and autonomous elements. However, as Raja Parasuraman predicted, improvements in autonomy and AI have not reduced the roles of the human; rather, they have changed them—sometimes dramatically (Parasuraman and Manzey 2010).

Over the last few decades the US Army Combat Capabilities Development Command Army Research Laboratory (CCDC ARL) and its collaborators have conducted research in AI/machine learning (ML), robotics systems, autonomous systems, natural language processing (NLP), human–robot interactions (HRI), and human–agent teaming (HAT) (Evans et al. 2018; Holder 2018; Barnes et al. 2019b). The objective of this report is to summarize the essential HRI research funded by the Human Research and Engineering Directorate (HRED) focusing on the changing role and supporting technology for the Soldier/operator. The summary will discuss lessons learned as well as design principles in an encapsulated form rather than conduct an extensive literature review. The format is topical, beginning with teleoperations of robots, then discussing topics such as multimodal control and display, adaptive systems, RoboLeader (a planning agent), models of trust and transparency, design implications of the human–agent interface, and a discussion of current and future efforts related to bidirectional communications between humans and agents. The arrangement of the discussion is meant to reflect how the scientific zeitgeist changed over the course of the HRI research program, from human control of robots to higher-level interactions with intelligent agents (Barnes and Evans 2010; Chen and Barnes 2014; Chen et al. 2018).

2. HRI and Teleoperations

Whereas teleoperations was an initial focus of HRI, it is still an important feature of robotic control because of the many fine grained uses of robots; for example, for finding and disarming improvised explosive devices (IED) and mines (Bodenhammer 2007), search and rescue (Khasawneh et al. 2019), and space-related applications (Guo et al. 2019), just to name a few. Indeed, Chen et al. (2007) concluded that, even for both autonomous and semi-autonomous systems, teleoperations will remain an important option as a backup during dangerous or unusual situations. They summarized the perceptual and cognitive issues for designing a teleoperated system (Table 1). For remote viewing, the world looks remarkably different; a limited field of view (FOV) causes the world to look constricted, making driving the robotic vehicle difficult especially because the sensor orientation and the position of the robot may be difficult to judge. The frame rate needs to be less than 10 Hz for optimal viewing (for a comprehensive review on frame rate, see Chen and Thropp 2007), and time lags over 170 ms degrade operator performance when driving the remote vehicle. The most effective frame of reference for the camera viewpoint is task dependent. The egocentric and exocentric viewpoints each have advantages and disadvantages, and integration of information from both may be difficult for the operator. Remote viewing may also cause motion sickness because of the conflicting vestibular and motion cues. For tasks such as disarming an IED, poor depth perception obtained from 2-D cameras can be ameliorated with stereovision (Bodenhammer 2007).

Table 1 Cognitive and perceptual issues for teleoperations (adapted from Chen et al. 2007)

	Issues	Suggestions
1	Limited field of view (FOV): erroneous speed and distance judgments, peripheral vision loss, degraded remote driving	Increase FOV or possibly use multiple FOVs for different tasks (multiple FOVs may be confusing – require training)
2	Robot sensor orientation and attitude of robot misperceived by operator	Track-up view for navigation; change to North-up view for map coordination tasks (e.g., recon tasks). Reference robot attitude in terms of gravity
3	Poor depth perception affecting size and distance for driving, manipulation, and navigation	Stereoscopic displays; caution: can cause motion sickness depending on the type of displays and individual susceptibility
4	Camera viewpoint and frame of reference: egocentric – do not see peripheral information; exocentric – loss of immediacy	Dual views and/or add peripheral cues for egocentric; consider specialized camera views
5	Slow frame degrades motion and spatial perception	10 Hz minimum; also augmented reality
6	Time delays from robot sensor to display more detrimental than delay from operator to robot	Minimum depends on task: 170 ms for driving; predictive displays help for longer lags
7	Attention switching for multiple sensors	Auditory and visual momentum cueing among views
8	Possible motion interference and motion sickness	Engineer interface to minimize vibratory and motion effects

ARL experiments at Fort Leonard Wood initially found stereovision difficult to use because an uneven depth of field (DOF) made the images fuzzy depending on the distance that the camera was from the target object. However, when DOF was sharpened over its range, mine detection and robot arm manipulation resulted in improved stereovision performance compared to 2-D viewing (Bodenhammer 2007). Edmondson et al. (2012) investigated pairing a haptic controller with the improved stereovision camera, showing better performance with the combined system versus baseline conditions (Fig. 1).

Fig. 1 Haptic controller and robot with a stereoscopic camera (Edmondson et al. 2012)

Pettitt et al. (2013) reported HRI design issues while comparing different types of autonomy with teleoperations during field exercises at Fort Benning. Semi-autonomy reduced workload but in some conditions the perceptual cues from the operator's camera-based teleoperation and the automated obstacle avoidance capability conflicted. According to Soldier feedback, the hardest thing to learn was how to maneuver the robot around objects when the obstacle avoidance system was being used. Although the obstacle avoidance system appeared to adversely impact course completion times, it did reduce the number of driving errors over those with teleoperation and it also reduced the percentage of times the operators had to stop driving in order to perform the secondary task compared to teleoperations. This capability needs to be refined so Soldiers can choose how to efficiently maneuver around the detected objects. This confusion between the manual and automated handoff emphasized that task allocation between operators and automation needs to be designed carefully to ensure that operators internalize their portion of the task so that it is aligned with the automated portion.

Telepresence is the use of robotic sensors to create the sensation of the operator being in the environment monitored by a remote robot. Elliott et al. (2012) in collaboration with the TNO (Toegepast Natuurwetenschappelijk Onderzoek) laboratory in the Netherlands, found that the TNO telepresence robot (stereo-vision and stereo-audio) improved performance in a field exercise at Fort Benning, Georgia. Although dismounted Soldier participants felt the robot and the interface equipment were too cumbersome in its current configuration, feedback for the overall concept was very positive. They particularly appreciated the immersive effect of head-controlled camera views and remote audio.

HRED-supported research at the University of Central Florida (UCF) compared manual control of robots with different types of automation in an urban simulation environment. They found that manual control was suboptimal for detecting possible IEDs from robotic vehicles traveling at moderate speeds. However, they found that humans were superior at identifying types of roadside objects and making tactical decisions, suggesting that a combination of manual control and various degrees of automation was better than either type of control alone (Barnes et al. 2014).

An important consideration for designing future systems such as the NGCV (RV) is the targeting function of the remote vehicle. Chen (2010) reported the results of four simulation experiments in which an operator either controlled or supervised a robotic vehicle from a manned platform, including conditions when the operator conducted gunner functions concurrently. Teleoperations of the robot degraded gunner functions but making the robotic vehicle autonomous degraded using the robot for surveillance (i.e., operator out of the loop). The results indicate that robot control and gunner functions should be performed by separate operators. Aided Target Recognition (AiTR) alerted the vehicle operator regarding potential targets and thus reduced the gunner's workload; however, the benefits of using the aid depended on individual differences. For participants who reported high attentional control, false-alarm prone aids interfered with performance more so than for conditions utilizing miss-prone AiTRs. Higher attentional control presumably allowed operators to attend to the target display while monitoring the aid but a high rate of false alarms caused them to ignore the AiTR alarms (disuse), even in conditions when it was accurate (Parasuraman and Riley 1997). The opposite effect (misuse) was found for participants who reported poorer attentional control; they tended to over-rely on miss-prone aids at the expense of attending to the target display. Individual differences also affected the utility of target alarms, leading Chen to conclude that remote operator interfaces and training should be designed to be flexible so that they can be adapted for individual differences.

The use of semi-autonomous controls such as way-point navigation can reduce workload while giving the operator greater control of the robotic vehicle's other functions. However, hybrid options with manual, semi-autonomous, and autonomous functionality promise greater flexibility in future combat environments, as long as the different modes are designed to be integrated in such a way to support the operator's mental model of the tasking environment. In particular (Pettit et al. 2013, Wright et al. 2018), Soldier supervision of autonomous and semi-autonomous systems must not be at the expense of their overall situation awareness (SA) (Chen and Barnes 2014). Interfaces for autonomous systems must consider potential overload when operators are controlling/supervising or monitoring autonomous systems in addition to their other combat tasks. The following section discusses HRED multimodal research, whose purpose is to distribute information over multiple sensory channels and develop naturalistic control devices to enable operators to focus on their mission objectives while maintaining SA.

3. Multimodal Intuitive Displays for Autonomous Systems

3.1 Introduction

There is no doubt that autonomous capabilities will be instantiated throughout military systems on land, sea, and air. Significant achievements in advanced autonomous capabilities are increasing in rate and scope (Martin et al. 2019), generating many issues for policy regarding use (Williams and Scharre 2015). Autonomous robotic systems augmented by AI are being designed to enhance sensor-based capabilities and information distribution. Emerging concepts in net-centric and asymmetrical warfare will rely on this capability to "push" information autonomously, in addition to being easily "pulled" from existing sources (Lacdan 2019).

Given this heightened capability for information push and increasing need to interact with multiple autonomous or semi-autonomous systems, along with the need for rapid action in dynamic context, reducing information overload has become a design driver for developing systems. Suboptimal decision making, slower response times, and generally poorer performance can result if operators are too focused on processing information rather than performing tasks (Wickens 2008a). Superior quality of information is not sufficient, nor is rapid delivery, rather it is essential that information be packaged in a format that can be quickly comprehended (Mitchell et al. 2004). The design challenge is to create interfaces that enable rapid understanding and more naturalistic human-systems interactions (Elliott and Redden 2013).

In many tactical situations, whether stationary or on the move, it is the visual channel that is overwhelmed with incoming information. For these situations, Wickens (2008b) established a key principle based on numerous studies stating that an overall reduction in cognitive loading can be obtained if information is distributed strategically over multiple sensory channels. The situation of the aircraft pilot was studied extensively by Wickens, who used speech/audio alerts to enhance comprehension (Wickens 2002). In turn, ARL researchers worked on audio and speech issues as they relate to Army combat scenarios. They have also collaborated with leading investigators at TNO Netherlands and the Army Aeromedical Research Laboratory (Fort Rucker, Alabama) to develop tactile options for Army

Warfighters in high workload situations (Elliott et al. 2007, 2010; Duistermaat et al. 2007). From the start, the focus has been on usability as well as multisensory application. For a display to be intuitive it must be easy to learn, utilize preexisting knowledge when possible, and be easy to use in an operational context.

The development of multisensory intuitive displays must first identify design factors such as operational context, existing workload, and information requirements, along with principles of multisensory display design. Ideally, display design would begin with cognitive task analyses (Crandall et al. 2006), to identify task demands, task goals, and information requirements, as well as to drive systematic application of multisensory principles. Distribution of information across sensory channels is more effective with tasks having a common goal (e.g., driving and listening to navigation information) than across competing goals (e.g., visual search of complex terrain while listening to radio transmission of unrelated information). Component tasks should be trained to automaticity when possible (e.g., basic driving/flying tasks should be accomplished with ease before adding more channels of communication). The next section will discuss findings related to each sensory channel (visual, audio, speech, tactile) and some effective combinations.

3.2 Visual Displays

The ubiquitous use of visual displays for maps, graphics, and camera-based information—whether in stationary settings such as command centers or during mobile operations—attests to the importance of the visual channel (Fig. 2). For command centers and vehicle cockpits, map-based displays are essential yet complex, often with layers of information filters leading to the possible overuse of visual channels unless the displays are augmented by other modalities. Incoming information can be distributed through speech and tactile means; in addition, alarms and alerts are vital for attention management. For the dismounted Soldier on the move, additional issues arise that further encourage the use of alternative information displays. One is the "heads-down" nature of a handheld visual display, which will distract the user from immediate surroundings. Other issues pertain to successful night operations, where use of visual displays can expose the user's location to the enemy. These considerations suggest careful consideration of alternate means of information display, where possible.

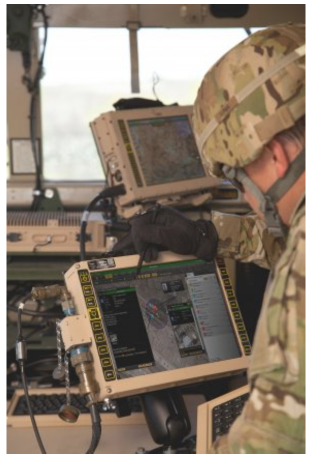

Fig. 2 Blue force tracking display within combat vehicle

Advantages of visual displays:

- Best for complex, map-based information; graphics; camera-based video.

Disadvantages of visual displays:

- Requires line of sight
- Often leads to information overload
- Often leads to attention tunneling
- Not covert during night-operations
- Higher workload to interpret direction cues
- Handheld options interfere with weapon use

3.3 Audio-based and Speech Displays

Audio Alerts

ARL research supports the use of audio cues for attention management, while attending to a visual display. ARL sponsored a meta-analysis of 24 studies comparing visual with visual-audio displays, that showed visual-audio multisensory displays were more effective than all-visual displays (Burke et al. 2006). The accumulation of findings provides strong support for reducing visual workload and attention management through use of audio alerts. Haas and van Erp (2014) detail many specific recommendations for the design of multimodal systems and they summarize the types of displays that favor the use of tactile or auditory cues.

Investigations of audio alerts have addressed perceptions of urgency (Haas and Casali 1995), for improved effectiveness of audio alerts. These principles generalized to applied endeavors, such as improving the design of the multiple sensor mine detection system (MDS) (Vause et al. 1999; Ferguson et al. 2000). Noting that MDS operators commonly suffered from undiagnosed hearing loss, they underscore the need for consideration of user characteristics in the design of signals. To avoid spectral regions of lower sensitivity, they advise that audio cues contain frequency components in multiple frequency ranges; at least one below 1 kHz and at least three between 1 kHz and 4 kHz, where hearing is most sensitive. The resulting signal design for the Hand-held Stand-off Mine Detection System (HSTAMIDS), shown in Fig. 3, incorporated a 500 Hz complex signal containing frequency content spanning several octaves.

Fig. 3 HSTAMIDS/3-D audio for communications

The operator should be able to learn the features of the auditory icons that discriminate between the types and sizes of mines easily. For example, they suggest that the pulse rate can be increased or decreased as a function of the size of the mine detected. Variation of other features such as tonal combinations, rhythmic patterns, and spectral effects (e.g., sharpness, brightness) can also distinguish between the types of mines detected. Use of robotic systems for mine detection would benefit directly from the HSTAMIDS research; 3-D audio conveys information about the spatial location of the sound source or uses binaural difference cues to control the perceived spatial locations of sounds within the acoustic environment. For example, by assigning each channel of one's radio communications to a spatially separate location, speech recognition benefits from the perceptual cues that allow the listener to selectively attend to only the desired channel, as demonstrated by Haas and her colleagues (Haas et al. 1997).

They demonstrated that 3-D audio, in this case, the spatial separation of radio streams, led to improved responses to radio communications, when three channels were present. In another study, spatially separating the source of radio communications input in a command and control vehicle improved speech intelligibility by as much as 15% and was the preferred option (Vause et al. 2000). Further, because sound location is perceived automatically, colocalizing auditory signals, such as alerts and warnings with the signal's source, or the item requiring attention, reduces the time required to respond. In a comparison of helicopter cockpit warning systems, the response time to visual signals augmented with 3-D auditory icons was significantly less than that of visual signals alone (Haas 1998).

ARL publications have summarized best practices for auditory signal design. Researchers from the Air Force Research Laboratory (AFRL) and ARL jointly published a guide to auditory displays (Letowski et al. 2001). Haas and Edworthy (2006) jointly authored a comprehensive book chapter on auditory signal design. A book on the design of helmet-mounted displays was the result of collaboration between ARL and the US Army Aeromedical Research Laboratory (Letowski et al. 2009). The book documents many of the basic psychophysical principles guiding audio and visual displays. The current version of MIL-STD-1472 section 5.3.1 (MIL-STD-1472G 2012, p. 118–130) summarizes many best practices with respect to the design of auditory warnings and signals and is the result of research conducted by ARL and other Department of Defense (DOD) organizations.

Speech

Speech is integral to Army operations. Radio-based communications are both sent and received, such that speech can be investigated as a display (receipt of information) and as a controller (sending of information commands). In this way, speech as a communication channel can be considered as another baseline, along with visual displays. The advantages of speech when added to visual display functions has been widely investigated in aircraft cockpit contexts (Wickens 2008a) to communicate critical information, improve attention management through alarms, and provide information status updates. These advantages can easily be generalized to ground vehicle and robotic contexts.

Speech has also been demonstrated to reduce workload when used in robot controller contexts, and is envisioned as integral to seamless interaction with autonomous assets. In an ARL-sponsored study performed at a robot maneuver course at Fort Benning, speech-based robot control was evaluated within the context of two visual display conditions (Pettitt et al. 2014). The overall effectiveness of the speech–visual combinations was significantly affected by the nature and referenced labeling of the visual information. ARL sponsored a series of investigations of speech-based commands, conducted by the University of Central Florida (Barber et al. 2014; Teo et al. 2014; Harris and Barber 2014). Studies were performed at the university and also with Soldiers at Fort Benning, Georgia. These commands were tailored to human–robot situations, in order to identify the most effective COTS speech recognition devices and develop an intuitive set of speech commands to convey commands such as direction, distance, and surveillance-reconnaissance tactics, from an operator to a robotic asset. These issues become more important as robotic assets become more autonomous and associated with more complex commands.

Advantages of speech and audio:

- Traditional Soldier–Soldier communication. Generalizes to Soldier interactions
- Lowers workload/increases comprehension in stationary settings
- Lowers workload when used to give commands to robotic assets
- 3-D audio cues can clarify multiple communication channels
- 3-D can provide direction cues
- Effective during poor visibility

Disadvantages of speech and audio:

- Not covert when silence required
- Not as effective in noisy context
- Front-back cone of confusion (3-D audio)

Speech has been found to be more effective when combined with another communication channel. In a study comparing effectiveness of speech and tactile cues for communicating direction and distance, researchers established that the combination of tactile direction cues and speech for distance information resulted in a faster, more accurate response than when both direction and distance were communicated with a single channel (Hartnett et al. 2018). Soldiers stood stationary and informed of the correct eight cardinal directions (i.e., north, northwest, west, southwest, south, southeast, east, northeast) for their location. They stood within a ring that was marked with reference points not pertaining to direction. They were given direction and distance cues using tactile and/or speech cues (see Fig. 4).

Fig. 4 Speech and tactile cueing for direction and distance information

3.4 Tactile Displays

There is increasing consideration and use of tactile displays for attention management, direction and spatial orientation information, and short communications. A particular advantage of tactile displays is the potential for pre-attentive processing, particularly for direction and spatial orientation cues (Elliott et al. 2014). Given the promising results regarding the use of tactile cueing for dismount Soldier navigation. ARL supported a multiyear program of research in tactile displays for direction cueing. Cognitive task analyses had identified Soldier navigation during movement to contact as one that was particularly high in visual workload (Mitchell et al. 2004).

ARL-sponsored meta-analyses of existing experiment-based comparisons provided foundational support for focused research on tactile displays. Meta-analyses of over 40 empirical studies showed significant improvement to performance outcomes when tactile cues were added to visual displays (Elliott et al. 2009). Field-based investigations first identified tactile cueing systems that were most effective during strenuous Soldier movement (Redden et al. 2006) and that could effectively convey short communications as well as direction information (Pettitt et al. 2006). Tactile cues were as easily comprehended as the arm and hand signals, while also having the advantage of working when out of line of sight and during poor visibility. Tactile cues have the advantage of relative stealth, and can remain effective when there is a need for operational silence, light security, or in conditions of poor visibility.

While many ARL studies have shown advantage of tactile cueing for direction and alerts (Krausman et al. 2007; Krausman and White 2008; White et al. 2012; Elliott et al. 2015, 2018, 2019a), the most dramatic demonstration of effectiveness was accomplished when the tactile belt was added to a chest-mounted visual display similar to the existing Nett Warrior concept, evaluated during night operations involving waypoint navigation and receipt of incoming messages. They used a standard chest-mounted visual display, consistent with Nett Warrior concepts, integrated with the tactile belt system. Results showed that missions performed with the tactile cueing were associated with reduced mission times, increased navigation accuracy, and lower reported experience of cognitive workload, effort, and frustration. Soldiers reported being more situationally aware of their surroundings and having better control of their weapon. In addition, when in the "tactile belt on" condition, Soldiers very rarely checked the visual display, averaging fewer than 2 times, compared to an average of over 17 times when the tactile guidance was not available. This was described by the Soldiers as the most critical operational advantage of the tactile guidance. Evaluation of tactile message comprehension was over 95%, when cues were presented only once during the waypoint navigation trials. Figure 4 shows a Soldier-participant in the night operations study, using the chest-mounted visual display. The torso tactile belt is worn under the uniform, over the T-shirt. It also shows the disadvantage of the visual display system when the participant is "heads-down". In contrast, the tactile belt system was described by the Soldier participants as "hands-free, eyes-free, and mind-free".

Tactile cueing has also proven effective for robot control. When added to robot control devices to provide direction cues to the operator using camera-based teleoperation, the operator could use a smaller robot control device as effectively as one with a larger video camera screen (Redden et al. 2009), because the tactile direction cueing minimized the need for constant reference to a map display.

More recently, tactile options have proven to be a viable means of human–robot bidirectional communications (Barber et al. 2013; Barber et al. 2015). Given that robotic assets are often out of line of sight, tactile communications offered the potential of covert communications from the robot to the controller, through short messages representing alerts or status updates. A series of experiments investigated the capacity of operators to learn and remember tactile cues. ARL sponsored laboratory-based investigations of tactile-based communications, led by Daniel Barber at UCF (Barber et al. 2014). Investigations centered on the development of tactile cue learning, based on a lexicon of tactile cues. Differences in learning and performance were also identified based on tactile cue characteristics; however, differences were ameliorated with refinements to the training process.

Studies of tactile messaging were also accomplished using Soldier-based evaluations conducted at Fort Benning, with messaging similar to traditional Army hand and arm signals and basic alerts that would be sent from robot to user. It was hypothesized that certain tactile cue characteristics (i.e., tempo, frequency complexity) would affect the perception of the tactile cue, in terms of tactile salience (i.e., the ease with which a tactile cue is perceived), which would in turn affect the ease of learning and recall. Soldier-participants were able to learn 12 different tactile cue commands in less than 30 min. They were able to accurately recall cue meanings several hours after the first training session (Elliott et al. 2019a; Elliott et al. 2019b). Tactile cue characteristics were significantly associated with levels of tactile salience, ease of learning, and accuracy of recall, leading to several guidelines for the design of multi-factor tactile cues for enhanced multimodal bidirectional communication among robot operators and assets.

Advantages of tactile displays:

- Fastest response to direction cues
- Intuitive portrayal of spatial orientation (e.g., helicopter pilot)
- No need for line of sight
- Effective during poor visibility (e.g., smoke, fog)
- Effective during night operations (e.g., maintains light security)
- Covert (can adjust cues for silent low-frequency communications)

Disadvantages of tactile displays:

- Limited vocabulary/syntax options

3.5 Gesture-based Controls

Gestures, in the form of hand and arm signals, have always been used for military communications. Advantages are practical, allowing for rapid and covert coordination of actions in dynamic context. For the field of gestural controls, the technological progress is rapid, distributed among many different approaches, and yields a huge number of relevant publications. An ARL-sponsored review of the literature relevant to gestural control of robotic assets (Elliott et al. 2016) summarized studies regarding camera-based systems, instrumented gloves, and handheld approaches to gesture-based controls. For the dismount Soldier, the instrumented glove approach was identified as most suitable. Advantages and disadvantages of each approach will differ based on operational task demands and context.

ARL researchers are investigating the use of gestures when integrated with complex visual displays (see Fig. 5). They first developed a light convolutional neural network as a gesture classification model to detect hand gestures in real-time, using a data set of 6,700 examples, based on eight one-handed gestures and six two-handed gestures (Hansberger 2019). The network was able to find archetypal features of each gesture and classify new data samples by scanning a real-time stream of joint rotations. They also investigated issues related to "gorilla arm syndrome", where fatigue can arise from prolonged hand and arm gesture activity. Gestures were systematically modified to use a supporting device that allowed natural resting positions, and thus reduce fatigue. (Hansberger et al. 2018). These findings inform an ongoing effort to develop an interface using input from voice, hand gestures, and eye gaze to interact with information in a virtual environment (Hansberger et al. 2019). As an example of integrated performance, the operator could use eye gaze for selection, with further instruction from speech or gesture, depending on operator choice. Future efforts in this area include a series of experiments that will examine the performance, engagement, and user experience levels that the multimodal system provides within the virtual environment.

Fig. 5 Instrumented glove concept for visual displays (left) and dismount robot control (right)

Field-based evaluations of prototype gesture-based robotic control systems have demonstrated advantages for wearable gesture control devices. An instrumented glove prototype was found effective for robot maneuvering (e.g., moving through robot obstacle course) and fine control of the robotic arm (Hartnett et al. 2018, see

Fig. 6). The assessment included demonstration of pointing gesture for direction information. The pointing gesture was demonstrated successfully for robot movement

Fig. 6 Chest-mounted visual display, with torso tactile belt (right) worn under uniform

An instrumented glove was used for Soldier-to-Soldier communication of direction information, when integrated with a tactile belt (Hartnett et al. 2018). Soldiers in different locations could easily refer to a specific location using a pointing gesture, with the direction reflected accurately on the tactile belt. For example, if locations A, B, and C were points on an equilateral triangle, Soldier 1 at location A used the instrumented glove to point to threat location C. Soldier 2 at location B received the direction information through the tactile belt, which directly “pointed” to threat location C, thus minimizing any confusion that could arise from speech communication.

A smartwatch form factor for speech and gesture-based controls was demonstrated successfully for control of robotic mules (see Fig. 7). Results showed that while both command methods were fairly accurate and the speech commands were particularly effective when not immersed in a high-noise environment, using a more typical headset with a directional, noise-canceling microphone potentially could yield even better results. Gesture recognition is another potential area for improvement. The gesture recognizer used the native inertial measurement unit on the smart watch, which was sensitive arm motion on some trails, and generally had too fine a window of recognition. Even so, with few exceptions, Soldiers successfully maneuvered the robot using all the gesture and speech commands. Based on Soldier feedback and observations, the smart interaction device (SID) was easy to learn and use for both gesture and speech commands. Soldiers gave positive feedback on the operational relevance of both speech and gesture as a means of intuitive control of robotic assets.

Fig. 7 SID for advanced human–robot interaction and Lockheed Martin's SMSS robotic mule

Advantages of gesture communications:

- Naturalistic means for direction cues (pointing)
- No need for line of sight (instrumented gloves)
- Can be consistent with naturalistic movements
- Already used for Soldier–Soldier communications, easily generalized to human–robot interaction
- Covert (silent)
- Effective under high noise

Disadvantages of gesture communications:

- Limited vocabulary/syntax options
- Gesture recognition options differ in reliability

The combination of gesture and speech holds high promise for multimodal systems for robot control. Speech-based controls have been developed with the goal of natural language interaction. At this time, purely speech-based controls face a core challenge regarding communication of spatial relationships and explicit directions (e.g., "go to the east side of the third building behind the church"), and pointing gestures are expected to help clarify speech-based localization information.

While benefits have been demonstrated, challenges remain with regard to effective integration of speech and gesture, particularly in multi-object environments in a 3-D world. These more complex scenarios represent an uncertain and unstructured problem space. Similar issues are faced as researchers strive to develop an interface that integrates head-mounted visual display, speech, and gesture controls for the commander situated in a moving command vehicle (Neely et al. 2004). Gestures are most naturally effective in situations of physical co-presence, where the robot and the operator can establish a joint visual understanding of the environment, with physical and directional referents. This is particularly true if robot recognition of gestures is dependent on a camera-based system.

3.6 Summary Discussion

There is now overwhelming evidence of the potential benefits of multisensory displays. While laboratory-based studies search for generalizable principles, they are complemented by many simulation and field-based studies that demonstrate advantages across numerous operational contexts. A review of these studies can reveal significant implications given a particular context, and implications will differ as task demands and context differ. The advantages of particular communications will be different for a dismount Soldier on the move during night operations than for crewmembers within a command and control center or vehicle cockpit. However, the approach to multimodal design will be similar, beginning with in-depth cognitive analysis of task demands and information requirements.

Certain principles will generalize across context. As an example, torso-based tactile cues are consistently best for direction cues, resulting in the fastest response time. However, tactile cues are not effective to communicate distance, as that would require the operator to "count" the number of tactile pulses, for distance information. Thus, the combination of tactile cueing for direction along with speech-based cueing of distance information (e.g., "200 meters"), should result in a faster, more accurate response across operational context. In the same way, speech-based controls should result in a faster, more accurate robot navigation compared to joystick control (e.g., "move forward and turn left at the second building on the right, go to the back of the building and scan for threat"). The addition of pointing gestures can further clarify speech commands, particularly when the robotic asset is situated at a different angle from the operator (e.g., "go that way 500 meters"). Longstanding principles of aircraft cockpit pilot performance should generalize directly to vehicle commander operations (e.g., use of speech to augment visual displays). It is this type of complementary interaction that is sought to bring laboratory-based findings to bear on multimodal system design, and thus enhance performance and reduce cognitive workload within Army operations.

Certain principles of multisensory design will also cluster within an operational context. While the optimal multisensory display configuration will depend on factors specific to operational mission context and individual operator preferences, general predictions can be made. For dismounted troops, multisensory options can alleviate disadvantages such as "heads-down" attentional tunneling with wearable options that are hands- and eyes-free. For mounted vehicle commanders, audio and tactile displays can manage attentional focus and offer intuitive direction cueing. For mounted robot controllers, tactile displays offer haptic messaging of critical alerts from robotic assets, along with direction cueing and haptic feedback of terrain (e.g., haptic steering).

While general principles can ameliorate workload in visually overloaded operations, individual operator preferences should always be considered through design of conformable and adaptive displays. Soldiers and research participants consistently provide feedback urging capabilities that can be adjusted to user preferences. They would like to be able to turn any particular display off, "lower the volume" of any particular communication source, and choose the controller option best suited to operations. This capacity to choose options also provides the opportunity for redundancy that itself offers advantages in terms of effectiveness (i.e., redundant messaging conveys urgency and accuracy) and operational risk reductions, if one channel were to fail. The following sections will discuss the implications of increased autonomy and humans teaming with intelligent software agents. Most of the discussion will involve the operator during mounted missions. However, future dismounted Soldiers will also partner with agents conducting routine combat patrols (Selkowitz et al. 2016; Chen et al. 2018). In both situations, human–agent interactions will not only partially unburden Soldiers but also require additional attentional resources (Parasuraman and Manzey 2010; Chen and Barnes 2014). Multi-sensory interfaces will help Soldiers maintain their own SA by partitioning incoming communications related to the immediate situation separately from communications emanating from the agent.

4. Human–Agent Teaming

Over the course of the HRI research, the human role evolved from control, to supervision, to collaboration (Parasuraman et al. 2000; Kelley and McGhee 2013; Chen et al. 2018). Systems that heretofore were manually controlled with automated components are evolving towards autonomy in the sense that these systems are able to be aware of their environment, react to change, and alter their COA when necessary to achieve their prescribed objectives (Russell and Norvig 2009). An intelligent agent (IA) in this context is an algorithm that performs specified tasks autonomously under the auspice of its human operator (Chen and Barnes 2014). Teaming relationships between agents and humans potentially combine human meta-knowledge with the agent's use of specialized algorithmic solutions to rapidly solve complex problems (Draper et al. 2018). To be successful, HAT require a collaborative relationship between the agent and its human partner and is predicated on mutual transparency and bidirectional communication (Chen et al. 2018). Partnership between human and agents presents problems as well as advantages: the two types of intelligence are not symmetrical (Kahneman 2011; Chen and Barnes 2014; Barnes et al. 2019a). The world model of agents is not only constrained by its software underpinning but also by its difficulty in adjusting to novel events and its limited ability to anticipate human information requirements in a dynamic environment (McNeese et al. 2017).

Early research on HAT was conducted by ARL's Robotic Collaborative Technology Alliance jointly by Penn State University and ARL scientists who developed an agent architecture to support human–agent teams. They combined a cognitive model, Recognition-Primed Decision-Making (RPD), with an agent architecture, Collaborative Agents for Simulating Teamwork (CAST), resulting in RPD-enabled CAST (RCAST). In a simulation experiment, the teaming relationship of RCAST and human participants (H-A) was compared to human–human teams (H-H) conducting a command and control mission (Fan et al. 2011; Chen and Barnes 2014). H-A teams using RCAST performed better than H-H teams; however, as the tasking environment became more complex, the H-A conditions started to degrade, possibly indicating RCAST brittleness in a changing environment.

In contrast, mixed initiative models allow for human–agent flexibility by enabling either the human or the agent to act depending on the situation. Adaptive and Adaptable systems are variants of mixed-initiative control in which humans always maintain decision authority, either directly or by specifying conditions under which the agent reacts automatically (Parasuraman et al. 2007; Chen and Barnes 2014; Barnes et al. 2015; Barnes et al. 2019b). Adaptive automation has a number of definitions; in this case; it is the triggering of an agent based on the state of the operator (increases in error rate, electroencephalography [EEG] fluctuations) or the state of the environment (multitask loading) (Chen and Barnes 2014). As part of HRED-supported research, Parasuraman et al. (2007) reported the results of three studies utilizing unmanned vehicles in which adaptive triggering of an aided target recognition agent based on individual operator error rate was superior to non-aided performance and to automation triggers averaged over multiple operators. The results showed improvements in workload, situation awareness, and change detection, suggesting that adaptive systems should be based on individual performance, again indicating the importance of individual differences for HRI. Adaptable systems (Parasuraman et al. 2007; Miller and Parasuraman 2017) are designed to give the operator more control by allowing the operator to decide when an agent is instantiated. The disadvantage is that it contributes to workload by requiring operators to decide when to trigger a specific agent during high-workload mission segments.

Initial forays into human–agent teaming (HAT) research exposed problems as well as provided solutions. While in many regards, HAT appeared to be an optimal solution—joining human ingenuity with computer processing capability to increase the human's scope of effectiveness and the team's overall efficiency—familiar human factor issues with automation surfaced—most notably human out-of-the-loop (and automation bias (Parasuraman and Riley 1997; Wright et al. 2017). The following sections are discussions on how these and other issues were explored as part of ARL's HRI program starting with our discussion of the RoboLeader agent research program.

5. RoboLeader: Agent Control of Multiple Systems

RoboLeader was one of the initial ARL HRI projects to investigate human–agent efficacy (Chen 2010). The RoboLeader agent was a surrogate supervisor that managed multiple robots (Chen et al. 2011). Diverse studies using the RoboLeader paradigm resulted in a better understanding of the agent's contributions to human decision-making during manned-unmanned operations (Chen and Barnes 2012a, 2012b; Wright et al. 2016, 2017, 2018). If events occurred requiring route changes or allocation of new tasks to the robots, the agent computed needed changes and issued new instructions, but human operators authorized their execution. The results of the studies indicated that agent error rate was not as important as type of error (false alarms vs. misses) but that RoboLeader aiding proved useful even for situations with imperfect agent reliability (Chen and Barnes 2102a, 2012b). The most ubiquitous finding of these and similar HRI studies was the importance of individual operator differences in determining effectiveness of HAT interactions (Chen and Barnes 2014; Chen et al. 2018).

More recent RoboLeader studies examined how the amount of information the operator was given about the experimental environment (supervising convoy operations; Fig. 8) interacted with the transparency of the agent's reasoning. The operator was better able to override the agent when it was incorrect if the reasoning behind the rerouting decision was explained. However, when the agent gave the operator additional information (staleness of agent's report), it actually hurt performance compared to the reasoning alone condition, suggesting that extraneous and/or ambiguous information can be harmful (Lee 2012). In the second study, the operator was given more information about the study environment and performed the same tasks with the same levels of agent reasoning transparency as in the first study. Access to the agent's reasoning did not have the same impact on performance as in the first study; however, the addition of extraneous information again hurt operator performance. Together, these studies indicate that transparency of the agent's reasoning is most helpful when the operator has limited knowledge of their task and/or environment, and ambiguous information is detrimental to operator performance (Wright et al. 2017).

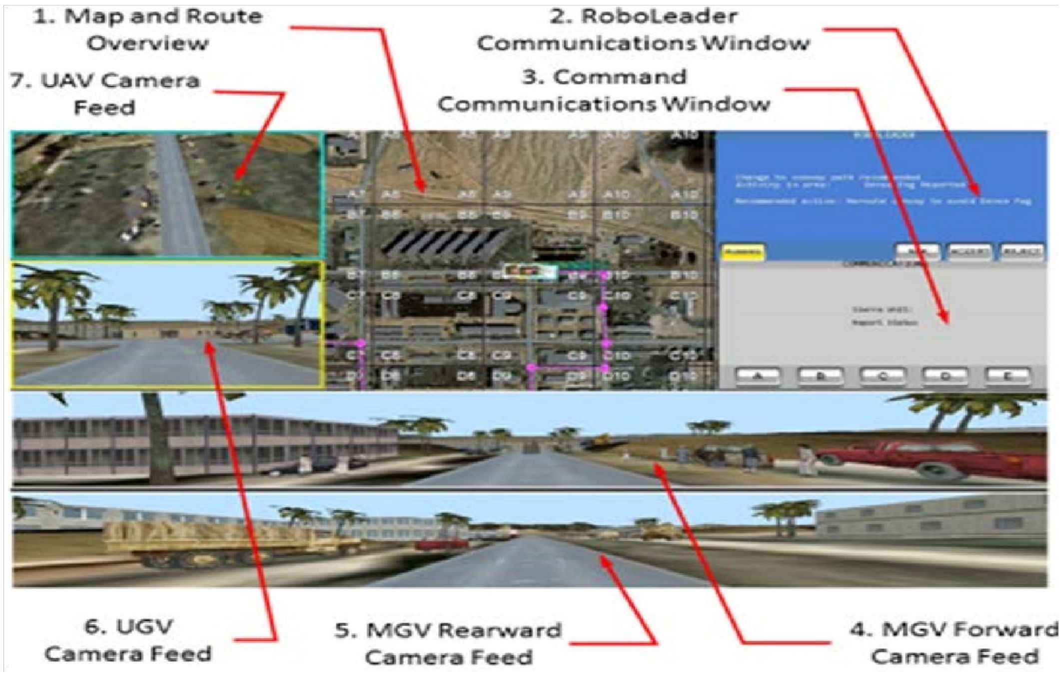

Fig. 8 The operator's control unit is the user interface for convoy management and 360° tasking environment. OCU windows are (clockwise from the upper center) map and route overview, RL communications window, command communications window, MGV's forward 180° camera feed, MGV's rearward 180° camera feed, UGV's forward camera feed, and UAV's camera feed (adapted from Wright et al. 2017).

6. Trust and Transparency

Lee and See's (2004) seminal paper on trust identified two essential components of HRI: trust and transparency. We discuss some of the research that their paper motivated including models of trust and transparency and supporting research that were developed as part of HRED's HRI program.

6.1 Trust Models

Calibrated trust is a requisite for collaboration among both human and human–agent teams to ensure that functionality is optimally distributed among team members. Calibrated trust was initially framed in terms of reducing automation misuse and disuse (Parasuraman and Riley 1997; Lee and See 2004; Lee 2012).

The human operator's misuse of automation was attributed to biases causing complacency such as ignoring information signaling automation failure (Parasuraman and Manzey 2010; Wright et al. 2016). Disuse is the opposite problem of not depending on automation when its use is appropriate. Human beliefs concerning the proper role of humans versus machine decision-making can result in automation disuse causing humans to ignore correct automated solutions (Beck et al. 2007). Dzindolet et al. (2003) found automation use sensitive to the lack of transparency; operator reliance waxed and waned depending more on their uncertainty concerning the reasons the automation was making errors than on the actual performance of the automated system. However, using a meta-analysis to investigate a large cross section of automation and robotic studies, UCF and ARL scientists found that the characteristics of the agent (especially its performance) was the main determiner of trust (Hancock et al. 2011). More recent analysis by Schaefer et al. (2015) found a significant (but moderate) effect on human-related factors as well as effects related to the characteristics of the agent. Hoff and Bashir (2015) developed a model of trust type based on an extensive review of the literature: disposition (depending on the person), situational (depending on the agent in a specific environment) and learned (depending on the person's experience with agents) suggesting that trust is related to individual differences and past experience rather than being a unitary process related to the agent. Schaefer et al. (2017) argue that situational distrust is partially caused by the operator not understanding the intent of the agent in complex environments. They developed a general model of trust (Fig. 9) indicating its relationship to human, agent (robot), and environmental factors as well as the human's past experience with the agent (Schaefer et al. 2019).

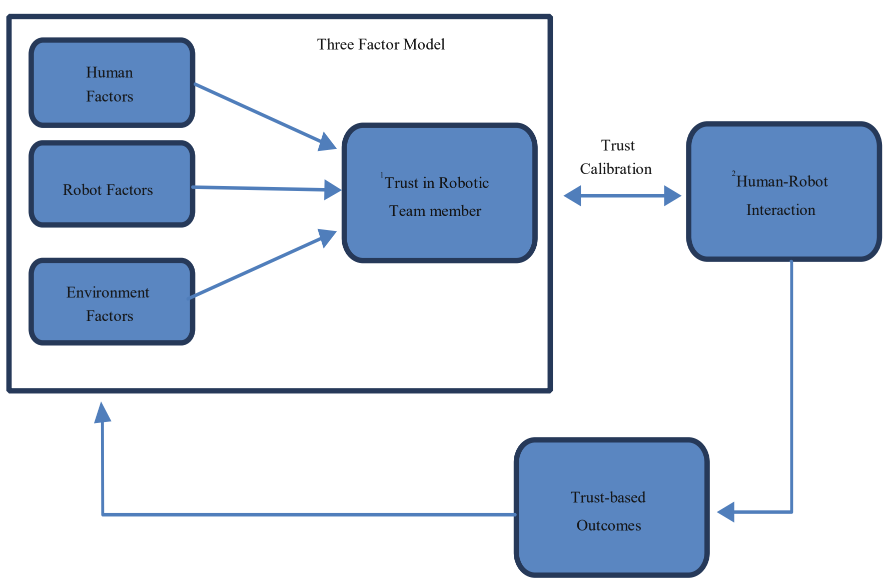

Fig. 9 Model of human–agent trust (Schaefer et al. 2019)

6.2 Situation awareness-based Agent Transparency (SAT) Model

Calibrated trust between humans as well as between humans and agents requires transparency (Dzindolet et al. 2003; Lee 2012; Chen and Barnes 2014; Lyons and Havig 2014; Schaefer et al. 2019). To address this issue, Chen et al. (2014) developed the Situation awareness-based Agent Transparency (SAT) model (Fig. 10) to delineate requirements for agent transparency. The SAT model was inspired not only by Endsley's (1995) SA model, but also the Beliefs, Desires and Intent model (Chen et al. 2018) and Lee's three P's: Purpose, Process and Performance (Lee and See 2004). By incorporating aspects of each of these models into a cohesive concept, researchers could now operationalize transparency to facilitate and organize research, as well as inform interface design and assessment. Similar to Endsley's (1995) model, SA subsumes 3 levels: Level 1–intent, plan, and action; Level 2–reasoning; Level 3–projected outcomes and uncertainty. However, in this case, SA refers to the agent's assessment of its SA in relation to its proposed COA (Chen et al. 2014). Studies encompassing multiple scenarios have shown that SAT information is incremental—each level adds useful information to the operator decision-making. In contrast, the utility of uncertainty information is mission specific. For large-scale planning, uncertainty information is generally useful because it alerts operators to potential limitations of the agent's plan. For squad levels missions, uncertainty information is not as useful because operators are time constrained and they require immediate "status at a glance" directions (Chen et al. 2018; Mercado et al. 2106; Stowers et al. 2016; Selkowitz et al. 2016).

Situation Awareness-based Agent Transparency

Level 1: Goals & Actions

Agent's current status/actions/plans

- Purpose: Desire (Goal selection)
- Process: Intentions (Planning/Execution); Progress
- Performance
- Perception (Environment/Teammates)

Level 2: Reasoning

Agent's reasoning process

- Reasoning process (Belief/Purpose)
- Motivations
    - Environmental & other constraints/affordances

Level 3: Projections

Agent's projections/predictions; uncertainty

- Projection of future outcomes
- Uncertainty and potential limitations; Likelihood of success/failure
- History of Performance

Fig. 10 SAT model (adapted from Chen et al. 2014)

7. Mutual Transparency

The success of HAT will ultimately depend on communications between the two entities. To reflect the importance of human–agent communications, the SAT model was recently expanded to encompass bidirectional transparency. Figure 11 illustrates mutual transparency of the human having SA of the agent and the bidirectional transparency of the agent's SA of the human (Lyons 2013; Chen et al. 2018). To support bidirectional understanding, HRED is collaborating with Institute for Creative Technologies (ICT) at the University of Southern California to delineate research necessary for human–agent communications (Wang et al. 2016; Pynadath et al. 2018; Barnes et al. 2019a). Besides mutual transparency, communications require a) media (e.g., graphics, voice, text, etc.); b) process (e.g., natural language processing [NLP]); c) AI underpinning of the agent such as machine learning (ML); and d) specialized instruments (eXplainable AI [XAI]) (Kelley and McGhee 2013; Barnes et al. 2019b). The agent must understand the human's requirements and vice versa, which for processes such as ML may require another level of translation to explain to the human the reasoning underlying the agent's solution. ML solutions such as reinforcement learning algorithms may be opaque due to the convergence on a solution based on induction rather than depending on deductive logic (Everett and Hutter 2018). ICT is developing XAI techniques to understand and parse ML solutions so they are transparent to its human team member (Chakraborty et al. 2017; Pynadath et al. 2018).

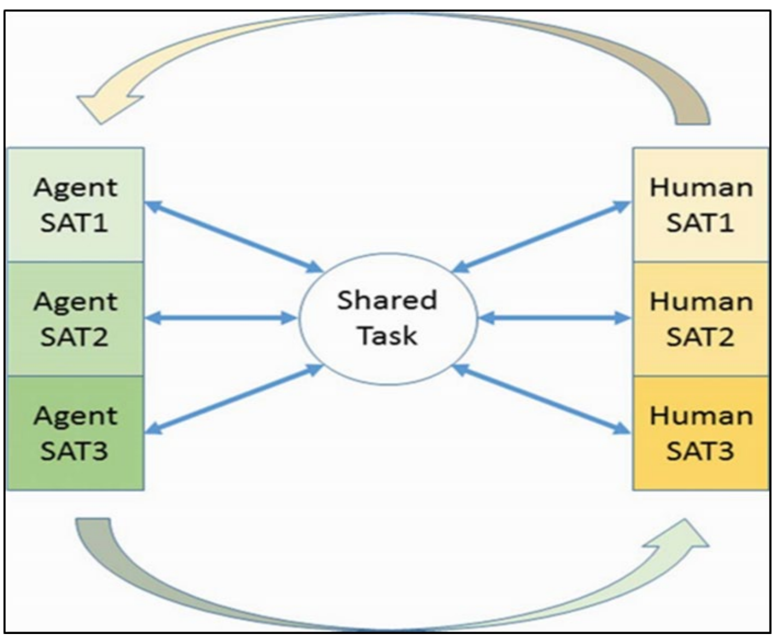

Fig. 11 Mutual transparency (Chen et al. 2018)

Mutual transparency is of particular import because human teams do not always interact through overt communications. Humans anticipate their partner's actions and information requirements reducing the necessity for overt communications (Cooke 2015). Salas et al. (2015) point out that "teams that communicate effectively may alternate between explicit communication, or overt transmission and acknowledgment of messages, and implicit communication, whereby information is more passively conveyed." One reason that humans anticipate rather than ask is because they have a *theory of mind* (TOM) that allows humans, especially those that have previous experience with their human team member, to have insight into the mental processes of their human partner (Astington and Edward 2010; Pedersen 2018; Mahey et al. 2014). ICT is conducting research on a recursive software model (*PsychSim*) to partially emulate a TOM for software agents as an initial step in enabling agents to develop shared mental models (SMM) with their human partner (Pynadath and Marsella 2005; Chen and Barnes 2014; Wang et al. 2016, 2018; Kwon 2018). Software SMM are in their early stages and fluid interactions between humans and agent still depend on overt communication techniques (Barnes et al. 2019b). However, HRI research on *Controlled English* shows promise by delimiting the size of the lexicons and focusing on specialized domains making human–agent communications practical for specific military missions such as civil affairs and military intelligence (Giammanco et al. 2015).

8. Transparency Visualizations for HAT

8.1 Autonomous Research Pilot Initiative (ARPI) SAT Visualization Research

Whereas we discussed SAT efficacy in various paradigms both in terms of increasing trust and in improving overall performance, the presentation of SAT-based information benefits from effective visualization techniques underlying the various implantations of SAT (Cha et al. 2019). HRED researchers supported two DOD joint programs to show the importance of the SAT framework as integral to the development of autonomous systems as part of the Autonomous Research Pilot Initiative (ARPI). They developed two visualization suites of displays based on SAT for 1) *Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies* (IMPACT) project in conjunction with the AFRL and the Navy (Mercado et al. 2016; Draper et al. 2018), and 2) *Autonomous Squad Member* (ASM) in conjunction with Navy researchers (Selkowitz et al. 2016; Wright et al. 2019). Both programs had multiple iterations and experimental tests to develop visualization concepts based on SAT enriched by concepts developed by the other services (Calhoun et al. 2018).

The HRED team designed transparency displays based on the SAT model and human factors requirements developed by the AFRL for the IMPACT planning paradigm (Calhoun et al. 2018). Figure 12 illustrates the final version of the interface portraying multiple levels of information for planning a complex littoral instillation defense mission showing autonomous assets chosen, optimal routes, projected outcomes, and uncertainties using both graphics and text annotations. Figure 11 compares SAT visualizations of two plans enabling the operator to choose between different solutions proposed by the planning agent. Each plan illustrates tradeoffs among time, coverage, fuel endurance, and general capability computed by the agent. The importance of each parameter is shown as the relative height on the individual bar graphs allowing easy comparisons between plan A and B. For example, for timeliness plan A is rated higher than plan B but for sensor coverage plan B is rated higher than plan A. The operator is able to choose the better plan based not only on the agent's suggestions but also on how the operator perceived the proposed COAs in relation to the commander's intent and updated mission information. Because the agent was not always perfectly accurate (e.g., due to constraints out of its control), incrementally increasing SAT information improved the operator's correct automation usage (Stowers et al. 2016). In summary, transparency visualizations enabled the operator to use his/her own knowledge of the ongoing mission as well as the SAT displays for effective trust calibration during military human–agent teaming missions (Mercado et al. 2016).

The ASM displays were developed as *SAT at-a-glance* information from an autonomous robot that supported an infantry squad engaged in a combat patrol (Fig. 13). The combat environment required rapid decision-making in seconds rather than minutes favoring graphical designs that were immediate rather than detailed (Selkowitz et al. 2016). Graphics indicting the robot's future trajectory (L1+L2+L3) improved the participants' SA whereas adding uncertainty cues (L1+L2+L3+U) did not significantly improve SA (Selkowitz et al. 2016).

Fig. 12 Improved transparency visualization for IMPACT experiments: direct comparison of plan options A and B (adapted from Stowers et al. 2016), in a more concise format, with uncertainty information in the text box. Updated mission information in the text box on the lower left-hand side.

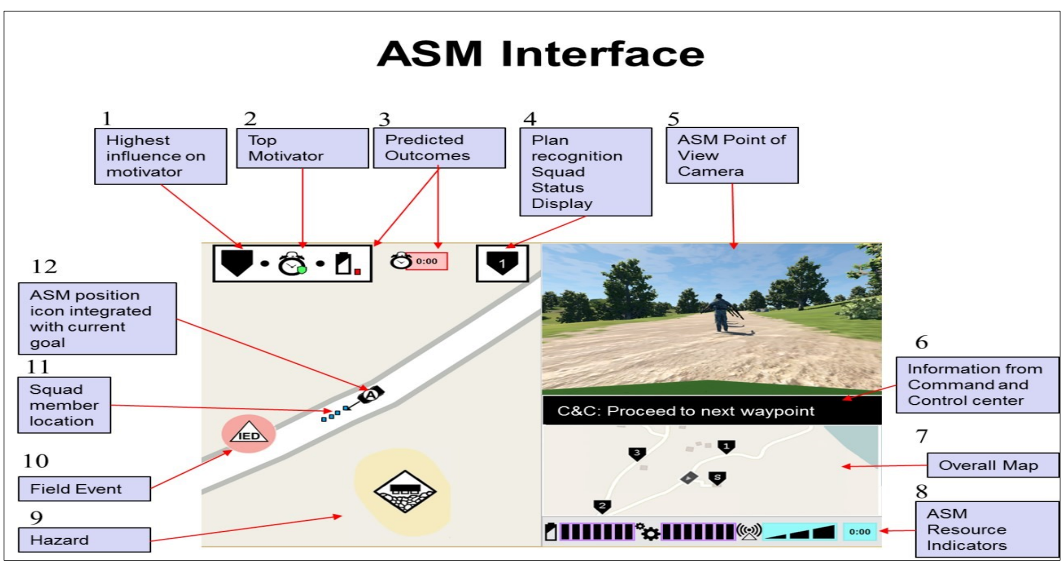

Fig. 13 Display for the ASM with annotations (Selkowitz et al. 2016)

8.2 Vehicular Displays

The military's fleets of UVs are becoming increasingly autonomous: eventually they will transition into robotic vehicles that are supervised rather than controlled.

The SA of the remote operator is maintained not only by feedback from smart agents but also by displays from the on-board sensors. The HRI program funded Israeli researchers from Ben Gurion University of the Negev who were evaluating displays for both Unmanned Aerial Systems (UAS) and ground vehicles (UGV) used for counter insurgency missions. Their initial studies in collaboration with HRED researchers at Fort Benning, Georgia, found Soldiers with 4-inch (diagonal) hand-held or 12-inch tablet displays could interpret intelligence information from remote sensors equally well but the helmet-mounted display conditions resulted in poorer performance, possibly because of binocular rivalry and eye-strain (Oron-Gilad 2014).

Later studies investigated remote operators supported by displays with multiple images so the operators were able to compare UAS and UGV views (Ophir-Arbelle et al. 2013). The imagery displays were able to combine scenes of the ongoing mission by integrating maps of areas of interest, UGV and UAS images. The interfaces improved performance over single imagery displays because remote operators were able visualize both the frontal (in the UGV) and planar (UAS) views of the ongoing mission within the context of the map display (Fig. 14).

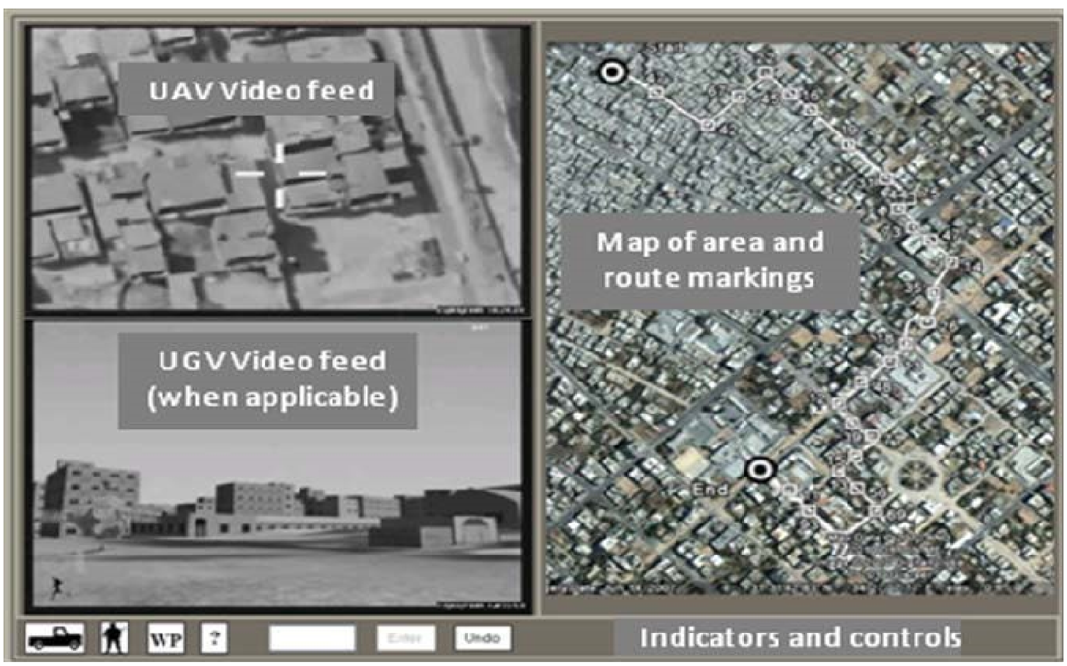

Fig. 14   Combined imagery views of UVs and map of area of interest (adapted from Ophir-Arbelle et al. 2013)

HRED researchers investigated display augmentations for Autonomous Navigation Systems (ANS) implemented in a robotic version of the Stryker combat vehicle. The augmentations (Fig. 15) displayed information extracted from the on-board ANS sensors to predict near-term and longer projected vehicle paths and obstacles

(Evans 2012). Evans investigated the aids in a field exercise at Camp Lejeune, North Carolina, showing that operator with the augmented visualizations relied on the ANS more so than in conditions where the augmented displays were not available. His study demonstrated that military operators are more likely to rely on (trust) autonomous options if they are shown projected outcomes.

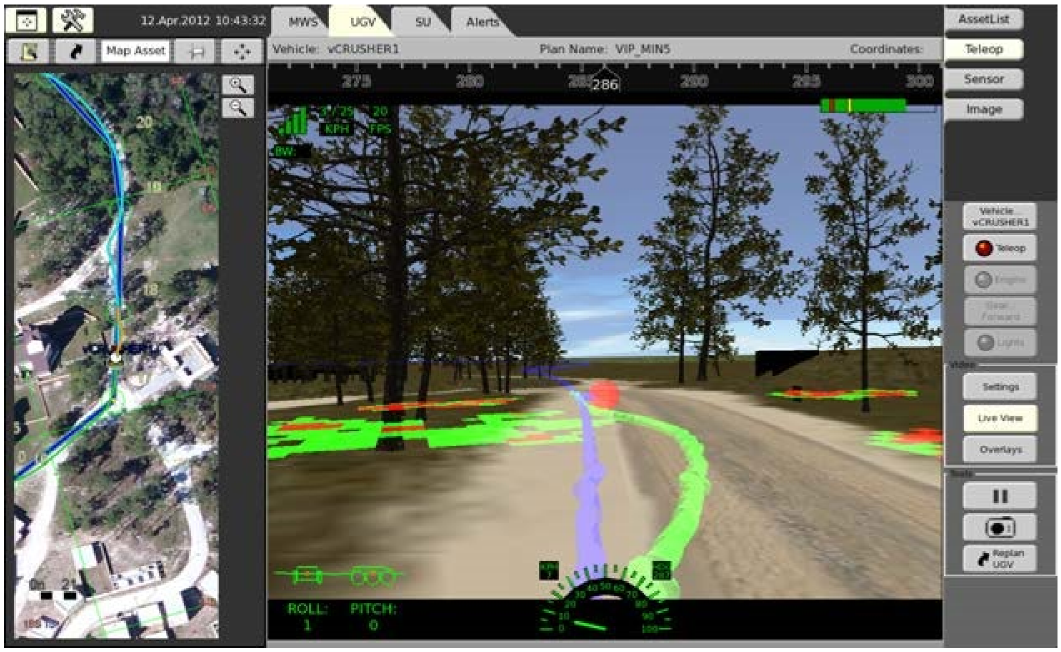

Fig. 15 Example of the Warfighter machine interface showing both the short-term (green) and long-term (blue) operator aids (Evans 2012)

8.3 Summary for Visualization

The lessons learned from these studies have shed light on operator performance and display design principles for complex military environments. When the amount of information to be processed seems overwhelming, humans will look to what they perceive as the "easiest" source of information, regardless of its appropriateness for the task at hand. As Wright et al. (2017) demonstrated, graphics are not always necessary, sometimes plain text will suffice. However, when comparisons are required and the amount of information to be processed is large, graphical visualization techniques have been shown to successfully convey information to the operator without increasing their cognitive workload (Wright et al. 2015; Mercado et al. 2016; Selkowitz et al. 2016). Displays that show both aerial and ground views of UV sensors within their geographical context will enhance the Soldier's overall SA of the battle space. Driver's trust of ANS increases if displays show both the near-term and longer-term projections of the vehicle's path.

In a more general review of visualization, Cha et al. (2019) conclude that "Successful techniques depend on re-creating the external reality of the environment in ways amenable to the human's mental representation of the processes involved." Specific paradigms have different requirements; visualizations contribute to understanding and trust when they intuitively capture the constraints and affordances in particular environments.

9. Conclusions

Combat will entail radically different technology from that which defined past military missions. Robotic aerial and ground vehicles, AI systems, augmented reality, autonomy, and Soldiers teaming with advanced algorithms will be utilized both by us and by our adversaries. An important principle underlying our research is that as robots and agents become more autonomous, Soldiers' taskings do not become easier—rather they change, often becoming more demanding. HRED's HRI program investigated the role of Soldiers in controlling/supervising/interacting with robots and agents during mounted and dismounted missions, most often in multitasking environments. Early research focused on teleoperations and the disadvantages of remote viewing and robot manipulation but we also reported mitigating technologies such as stereoscopic viewing, haptic manipulations, and multimodal solutions.

Multimodal interfaces were investigated to ameliorate the increased complexity of dismounted and mounted Soldiers controlling and eventually supervising robotic systems. The utility of multi-modal displays and controls is based on well-established cognitive research that has shown performance and SA gains by distributing cognitive resources over multiple modalities (visual, auditory, tactile, gesture). Based on our research findings, we discuss the performance advantages and limitations of each modality stressing the advantages of multimodal synergy in high workload or noisy environments.

Design of multimodal interfaces depends on the tasking environment and whether the Soldier is mounted or dismounted. Certain applications subsume both Soldier missions (e.g., tactile cueing for direction [left] and voice for distance [200 m]) whereas others are more mission specific. Gestures will be particularly useful for augmented reality applications because it is a natural way to transverse 3-D space and gestures are efficacious for small robot control as well. We also discussed the advantages of visual displays that showed different views of the battlespace and the usefulness of predictor displays for autonomous navigation.

As part of the RoboLeader program, we investigated agent supervision of multiple robots and manned vehicles while varying number of vehicles, type of agent errors, amount of automation, task type, operator differences, and cognitive biases. We concluded that using an IA acting as an intermediate supervisor reduced workload and improved performance even for agents with less than perfect reliability. We discussed the importance of calibrated trust and transparency for human–agent collaboration. The SAT model was developed to enable operators to understand the agent's current actions and intended plans, the logic behind the current actions and plans, and the projected outcomes and uncertainties. We developed SAT visualization concepts and principles for two DOD autonomy research programs (IMPACT and ASM) showing their efficacy for improving operator performance and for calibrating trust under diverse combat scenarios.

For human–agent teaming paradigms, we conclude that SAT information needs to be two directional—not only the operator's SA of the agent's intentions but also the agent's SA of human intentions. This is predicated on human-agent communication but it also implies an SMM to enable implicit interactions as well as mutual trust. Our future research objectives encompass more natural and more human-like interactions with agents (and embedded agents such as robots) through improved NLP strategies, software emulation of SMM, transparent ML, calibrated trust, and naturalistic multi-modal interfaces.

10. References

Astington JW, Edward MJ. The development of a theory of mind in early childhood. Montreal (Canada): Encyclopedia of early childhood; 2010. p. 1–5.

Barber D, Reinerman-Jones L, Matthews G. Toward a tactile language for human–robot interaction: two studies of tacton learning and performance. Hum Factors. 2015;57(3):471–490.

Barber D, Lackey S, Reinerman-Jones L, Hudson I. Visual and tactile interfaces for bi-directional human–robot communication. SPIE Defense, Security, and Sensing. 2013:87410U.

Barber D, Whleber R, Parchment A, Jentsch F, Elliott L. Development of a squad level vocabulary for human–robot interaction. Proceedings of the International Conference for Human Robot Interaction; 2014; Bielefeld, Germany. Springer-Verlag.

Barnes MJ, Chen JYC. Intelligent agents as supervisory assets for multiple uninhabited systems: RoboLeader. In: van Breda L, editor. Supervisory control of multiple uninhabited systems methodologies and enabling human–robot interface technologies. Neuilly sur Seine (France): North Atlantic Treaty Organisation; 2012. Report No.: RTO-TR-HFM-170.

Barnes MJ, Ning W, Pynadath DV, Chen JYC. Human-agent bi-directional communications; Forthcoming 2019a.

Barnes MJ, Evans AW, III. Soldier–robot teams in future battlefields: an overview. In: Barnes MJ, Jentsch F, editors. Human–robot interactions in future military operation. Surrey (England): Ashgate; 2010. p. 9–29.

Barnes MJ, Chen JYC, Jentsch F. Designing for mixed-initiative interactions between humans and autonomous systems in complex environments. Proceedings of 2015 IEEE Conference on Systems, Man, and Cybernetics (SMC); 2015 Oct; Hong Kong. p. 1386–1390.

Barnes MJ, Chen JYC, Jenstch F, Oron-Gilad T, Redden E, Elliott L, Evans A. Designing for humans in autonomous systems: military applications. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2014 Jan. Report No.: ARL-TR-6786.

Barnes MJ, Lakhmani S, Holder E, Chen JYC. Issues in human–agent communication. Aberdeen Proving Ground (MD) Army Research Laboratory (US); 2019b Feb. Report No.: ARL-TR-8636.

Beck H, Dzindolet M, Pierce L. Automation usage decisions: controlling intent and appraisal errors in a target detection task. Hum Factors J Hum Factors Ergon Soc. 2007;49(3):429–437.

Bodenhammer A. Assessment of stereoscopic displays for assisting in route clearing operations. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2007 Feb. Report No.: ARL-TR-4195.

Burke JL, Prewett M, Gray A, Yang L, Stilson FRL, Redden E, Elliott LR, Coovert MD. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. Proceedings of the International Conference on Multimodal Interfaces; 2006; Alberta, Canada. p. 108–117.

Calhoun GL, Ruff H, Behymer KJ, Frost EM. Human–autonomy teaming interface design considerations for multi-unmanned vehicle control. TIES. 2018;19(3):321–352.

Cha J, Barnes MJ, Chen JYC. Visualization techniques for transparent human–agent interface design. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2019. Report No.: ARL-TR-8674.

Chakraborty S, Tomsett R, Raghavendra R, Harborne D, Alzantot M, Cerutti F, Srivastava M, Preece A, Julier S, Rao R, et al. Interpretability of deep learning models: a survey of results. Presented at: IEEE Smart World Congress 2017 Workshop: DAIS 2017—Workshop on Distributed Analytics InfraStructure and Algorithms for Multi-Organization Federations; 2017 Aug 7–8; San Francisco, CA.

Chen JYC. Robotic performance in a multi-tasking environment. In: Barnes MJ, Jentsch F, editors. Human-robot interactions in future military operations. Surrey (England): Ashgate; 2010. p. 293–314.

Chen JYC, Barnes MJ, Harper-Sciarini M. Supervisory control of multiple robots: human performance issues and user interface design. IEEE Transactions on System, Man, and Cybernetics—Part C: Applications and Reviews. 2011;41(4):435–454.

Chen JYC, Barnes MJ. Human–agent teaming for multirobot control: a review of human factors issues. IEEE T Hum Mach Syst. 2014;44(1):13–29.

Chen, JYC, Barnes MJ. Supervisory control of multiple robots in dynamic tasking environments. Ergonomics. 2012a;55:1043–1058.

Chen, JYC, Barnes MJ. Supervisory control of multiple robots: effects of imperfect automation and individual differences. Hum Factors. 2012b;54(2):157–174.

Chen JYC, Barnes MJ. Agent transparency for human–agent teaming effectiveness. Proceedings of the IEEE conference on Systems, Man, and Cybernetics (SMC); 2015 Oct; Hong Kong.

Chen JYC, Haas EC, Barnes MJ. Human performance issues and user interface design for teleoperated robots. IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews. 2007;37:1231–1245.

Chen JY, Lakhmani SG, Stowers K, Selkowitz AR, Wright JL, Barnes M. Situation awareness-based agent transparency and human–autonomy teaming effectiveness. Theoretical Issues in Ergonomics Science. 2018;19(3):259–282. doi: 10.1080/1463922X.2017. 1315750.

Chen JYC, Procci K, Boyce M, Wright J, Garcia A, Barnes MJ. Situation awareness-based agent transparency. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2014 Apr. Report No.: ARL-TR-6905.

Chen JYC, Thropp JE. Review of low frame rate effects on human performance. IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans. 2007;37(6):1063–1076.

Cooke NJ. Team cognition as interaction. CDPS. 2015;24(6):415–419.

Crandall B, Klein G, Hoffman R. Working minds: a practitioner's guide to cognitive task analysis. Cambridge (MA): MIT Press; 2006.

Defense Science Board. Defense Science Board summer study on autonomy. Washington (DC): Under Secretary of Defense; 2016.

Draper M, Rowe A, Douglass S, Calhoun G, Spriggs S, Kingston D, et al. Realizing autonomy via intelligent adaptive hybrid control: adaptable autonomy for achieving UxV RSTA team decision superiority (also known as Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies [IMPACT]). Dayton (OH): Air Force Research Laboratory (US); 2018. Report No.: AFRL-RH-WP-TR-2018-0005.

Duistermaat M, Elliott L, van Erp J, Redden E. Tactile land navigation of dismount soldiers. In: de Waard D, Hockey G, Nickel P, Brookhuis K, editors. Human factors issues in complex environment performance. Europe chapter of the Human Factors and Ergonomics Society. Herzogenrath (Germany): Shaker Publishing; 2007.

Dzindolet M, Peterson R, Pomranky R, Pierce L, Beck H. The role of trust in automation reliance. International Journal of Human-Computer Studies. 2003;58(6):697–718.

Edmondson R, Light K, Bodenhamer A, Bosscher P, Wilkinson L. Enhanced operator perception through 3D vision and haptic feedback. Proceedings of SPIE, Unmanned Systems Technology XIV; 2012; Baltimore, MD.

Elliott LR, Jensen C, Redden ES, Pettit E. Robotic telepresence: perception, performance and user experience. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2012. Report No.: ARL-TR-5928.

Elliott LR, Mortimer BJ, Miller L. Designing tactile cues: factors affecting salience and recognition. Submitted to IEEE Transactions on Haptics. Forthcoming 2019a.

Elliott LR, Mortimer BJ, Pomranky-Hartnett R, Rapozo F, Pettitt R, Wooldridge R, Rapozomeran A, Mort G. Tactile cues: taction characteristics, salience, ease of learning, and recall. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2019b. Report No.: ARL-TR-8650.

Elliott LR, Mortimer BJ, Pettitt RA, Wooldridge RE. Assessment of wearable tactile system: perception, learning, and recall. Augmented Cognition: Users and Contexts. In: Lecture Notes in Computer Science, volume 10916. Las Vegas (NV): Springer; 2018. p. 67–77.

Elliott L, Mortimer B, Hartnett-Pomranky G, Zets G, Mort G. Augmenting soldier situation awareness and navigation through tactile cueing. Proceedings of the 17th International Conference on Human–Computer Interaction; 2015; Los Angeles, CA. p. 345–353.

Elliott L, Skinner A, Pettitt R, Vice J, Walker A. Utilizing glove-based gestures and a tactile vest display for covert communications and robot control. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2014. Report No.: ARL-TR-6971.

Elliott L, Redden E. Reducing workload: a multisensory approach. In: Savage-Knepshield P, editor. Designing soldier systems: current issues in Human Factors. Surrey (England): Ashgate; 2013.

Elliott LR, Coovert MD, Redden ES. Overview of meta-analyses investigating vibrotactile versus visual display options. Proceedings of the 13th International Conference on Human–Computer Interaction, Part II: Novel Interaction Methods and Techniques; 2009; San Diego, CA. p. 435–443.

Elliott L, Hill S, Barnes M. Gesture-based controls for robots: overview and implications for use by soldiers. Aberdeen Proving Ground, (MD): Army Research Laboratory (US); 2016. Report No.: ARL-TR-7715.

Elliott L, van Erp JBF, Redden E, Duistermaat M. Field-based validation of a tactile navigation device. IEEE Transactions on Haptics. 2010;3(2):78–87.

Elliott LR, Duistermaat M, Redden E, van Erp J. Multimodal guidance for land navigation. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2007. Report No.: ARL-TR-4295.

Endsley MR. Toward a theory of situation awareness in dynamic systems. Hum Factors. 1995;37:32–64.

Evans AW. Investigating the usefulness of operator aids for autonomous unmanned ground vehicles performing reconnaissance tasks. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2012. Report No.: ARL-TR-6651.

Evans A, Marge M, Stump E, Warnell G, Conroy J, Summers-Stay D, et al. The future of human robot teams in the Army: factors affecting a model of human system dialogue towards greater team collaboration. In: Savage-Knepshield P, Chen J, editors. Advances in human factors in robots and unmanned systems. Cham, Switzerland: Springer Open; 2017. p. 197–209. (Advances in Intelligent Systems and Computing, vol. 499).

Everitt T, Hutter M. Universal artificial intelligence. In: Abbass M, Scholz J, editors. Foundations of trusted autonomy. Cham (Switzerland): Springer Open; 2018. p. 15–47.

Fan X, McNeese M, Sun B, Hanratty T, Allender L, Yen J. Human–agent collaboration for time stressed multi-context decision making. IEEE Trans Syst Man Cybern A Syst Humans. 2011;40(2):306–320.

Ferguson L, Mermagen T, Vause N, Letowski T. Auditory signals for enhanced operator performance with hand-held mine detectors. Proceedings of the Detection and Remediation Technologies for Mines and Minelike Targets V. SPIE 4038; 2000; Bellingham, WA. p. 130–135.

Giammanco C, Mott M, McGowan R. Controlled English for critical thinking about the civil-military domain. Proceedings of the International Technology Alliance in Network and Information Sciences Annual Fall Meeting; c2015; Los Angeles, CA.

Goodrich MA, Schultz AC. Human–robot interaction: a survey. Found. Trends Hum Comput Interact. 2007;1(3):203–275.

Guo J, Liu Y, Kong X, Zhou S, Yang J. Analysis of key cognitive factors in space teleoperation task. In: Harris D, editor. Engineering psychology and cognitive ergonomics. HCII 2019. Lecture Notes in Computer Science, vol 11571. Springer, Cham; 2019.

Haas EC. Can 3-D auditory warnings anhance helicopter cockpit safety? Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting; 1998 Oct 5–9; Chicago, IL. p. 1117–1121.

Haas EC, Casali JG. Perceived urgency and response time to multi-tone and frequency-modulated warning signals in broadband noise. Ergonomics. 1995;38:2313-2326.

Haas EC, van Erp JB. Multimodal warnings to enhance risk communication and safety. Safety Science. 2014;29-35.

Haas EC, Edworthy J. An introduction to auditory warnings and alarms. In: Wogalter M, editor. Handbook of warnings. Mahwah (NJ): Lawrence Erlbaum and Associates; 2006. p. 189–198.

Haas E, Gainer C, Wightman D, Couch M, Shilling R. Enhancing system safety with 3-D audio displays. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 41, no. 2; 1997 Oct; Albuquerque, NM. p. 868–872.

Hansberger J. Real-time gesture recognition using 3D sensory data and a light convolutional neural network. ACMMM 19; 2019 Oct 21–25; Nice, France.

Hansberger J, Peng C, Mathis S, Shanthakumar V, Meacham S, Cao L, Blakely V. Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions. International conference on virtual, augmented, and mixed reality, VAMR 2017. Lecture Notes in Computer Science, vol 10280. Springer, Cham; 2017. p. 505–520.

Hansberger J, Peng C, Blakely V, Meacham S, Cao L, Diliberti N. A multimodal interface for virtual information environments. 2019 Human Computer Interaction International Conference; 2019 July 26–31; Orlando, FL.

Harris J, Barber D. Speech and gesture interfaces for squad level human robot teaming. Proceedings of SPIE Defense, Security, and Sensing – Unmanned Systems Technology; 2014 May 5–9; Baltimore, MD.

Hartnett G, Elliott L, Baraniecki L, Skinner A, Riddle K, Pettitt R. Wearable concepts for gesture-based robot control, direction localization, and communication to a tactile belt display. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2018. Report No.: ARL-TR-8597.

Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser E, Parasuraman R. A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors. 2011;53(5):517–527.

Hoff K, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors. 2015;57(3):407–434.

Holder E. Defining soldier intent in soldier-natural language processing interaction. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2018 Apr. Report No.: ARL-TR-8919.

Kahneman D. Thinking, fast and slow. New York (NY): Farrar, Straus & Giroux; 2011. p. 499.

Kelley TD, McGhee S. Combining metric episodes with semantic event concepts within the symbolic and sub-symbolic robotics intelligence control system (SS-RICS). SPIE Defense, Security, and Sensing; 2013 May; Anaheim, CA. p. 87560L.

Khasawneh A, Rogers H, Bertrand J, Madathil K, Gramopadhye A. Human adaptation to latency in teleoperated multi-robot human-agent search and rescue teams. Automat Constr. 2019;99:265–277.

Kwon D. Machines that learn like children provide deep insight into how the mind and body interact together to bootstrap knowledge and skills. Sci Am. 2018 Mar; p. 27–32.

Krausman A, White T. Detection and localization of vibrotactile signals in moving vehicles. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2008. Report No.: ARL-TR-4463.

Krausman A, Pettitt R, Elliott, L. Effects of redundant alerts on platoon leader performance and decision making. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2007. Report No.: ARL-TR-3999.

Lacdan J. Warfare in megacities: a new frontier in military operations. 2019 Jan 4 [accessed 2019 Aug 19]. https://www.army.mil/article/215752/warfare_in_megacities_a_new_frontier_in_military_operations.

Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors. 2004;46(1):50–80.

Lee JD. Trust, trustworthiness, and trustability. Presentation at the Workshop on Human Machine Trust for Robust Autonomous Systems; 2012; Ocala, Florida.

Letowski T, Karsh R, Vause N, Shilling R, Ballas J, Brungart D, McKinley R. Human factors military lexicon: auditory displays. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2001. Report No.: ARL-TR-2526.

Letowski T, Schmeisser E, Russo M, Rash C. Helmet-mounted displays: sensation, perception, and cognition issues. Fort Rucker (AL): Army Aeromedical Research Laboratory (US); 2009.

Lyons JB. Being transparent about transparency: a model for human–robot interaction. In: Sofge D, Kruijff GJ, Lawless WF, editors. Trust and Autonomous Systems: Papers from the AAAI Spring Symposium (Technical Report SS-13-07). p. 48–53. Menlo Park (CA): AAAI Press; 2013.

Lyons JB, Havig PR. Transparency in a human-machine context: interface approaches for fostering shared awareness/intent. In: Shumaker R, Lackey S, editors. Virtual, augmented and mixed reality: designing and developing virtual and augmented environments. Proceedings of the International Conference on Virtual, Augmented and Mixed Reality; 2014; Crete, Greece. Cham (Switzerland): Springer International Publishing. p. 181–190. doi: 10.1007/9783-319-07458-0_18.

Mahey CEV, Moses LJ, Pfiefer JH. How and why: theory of mind in the brain. Dev Cogn Neuros 9. 2014;68–81.

Martin B, Tarraf DC, Whitmore TC, DeWeese J, Kenney C, Schmid J, DeLuca P. Advancing autonomous systems: an analysis of current and future technology for unmanned maritime vehicles. Santa Monica (CA): RAND Corporation; 2019 [accessed 2019 Aug 19]. https://www.rand.org/pubs/research_reports/RR2751.html.

McNeese N, Mustifa D, Cooke N, Myers C. Teaming with a synthetic teammate: insights into human-autonomy teaming. Hum Factors. 2017;60(2):262–273.

Mercado J, Rupp M, Chen J, Barber D, Procci K, Barnes M. Intelligent agent transparency in human–agent teaming for multi-UxV management. Hum Factors. 2016;58(3):401–415.

Miller C, Parasuraman R. Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. Hum Factors 2007;49:57–75.

MIL-STD-1472G. Design criteria standard: human engineering. Washington (DC): Department of Defense; 2012.

Mitchell D, Samms C, Glumm M, Krausman A, Brelsford M, Garrett L. Improved performance research integration tool (IMPRINT) model analyses in support of the situational understanding as an enabler for unit of action maneuver team Soldiers science and technology objective (STO) in support of future combat systems (FCS). Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2004. Report No.: ARL-TR-3405.

Neely H, Belvin R, Fox J, Daily M. Multimodal interaction techniques for situational awareness and command of robotic combat entities. Proceedings of the 2004 IEEE Aerospace Conference; 2004 Mar 6–13; Big Sky, MT. Piscataway (NJ): IEEE; c2004. p. 3297.

Ophir-Arbelle R, Oron-Gilad T, Borowsky A, Parmet Y. Is more information better? How dismounted soldiers use video feed from unmanned vehicles: allocation and information extraction considerations. Journal of Cognitive Engineering and Decision Making 2013;7(1):26–48.

Oron-Gilad T. Scalable interfaces for operator control units: common display to conduct MOUT operations with multiple video feeds (final research report). Be'er Sheva (Israel): Ben Gurion University of the Negev; 2014.

Parasuraman R, Riley V. Humans and automation: use, misuse, disuse, abuse. Hum Factors. 1997;39:230–253.

Parasuraman R, Sheridan TB, Wickens CD. A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans. 2000;30(3):286–297.

Parasuraman R, Manzey DH. Complacency and bias in human use of automation: an attentional integration. Hum Factors. 2010;52(3):381–410.

Parasuraman R, Barnes M, Cosenzo K. Adaptive automation for human robot teaming in future command and control systems. International Journal of Command and Control. 2007;1(2):43–68.

Pedersen T. Theory of mind. Newburyport (MA): Psych Central; 2018 Dec 10 [accessed 2018 July 18]. https://psychcentral.com/encyclopedia/theory-of-mind.

Pettitt RA, Redden ES, Carstens CB, Hooper D. Scalability of robotic controllers: effects of progressive autonomy on intelligence, surveillance, and reconnaissance robotic tasks. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2013 Jan. Report No.: ARL-TR-6182.

Pettitt R, Redden E, Carstens, C. Comparison of Army hand and arm signals to a covert tactile communication system in a dynamic environment. Aberdeen

Proving Ground (MD): Army Research Laboratory (US); 2006. Report No.: ARL-TR-3838.

Pettitt R, Carstens C, Elliott L. Speech-based robotic control for dismounted soldiers: Evaluation of visual display options. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2014. Report No.: ARL-TR-8411.

Pynadath DV, Marsella SC. PsychSim: Modeling theory of mind with decision theoretic agents. Proceedings of the International Joint Conference on Artificial Intelligence; 2005. p. 1181–1186.

Pynadath DV, Barnes MJ, Wang N, Chen JYC. Transparency communication for machine learning in human-automation interaction. In: Zhou J, Chen F, editors. Human and machine learning. Human–computer interaction series. Springer, Cham; 2018.

Redden ES, Carstens CB, Turner DD, Elliott LR. Localization of tactile signals as a function of tactor operating characteristics. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2006. Report No.: ARL-TR-3971.

Redden ES, Elliott LR, Pettitt RA, Carstens CB. A tactile option to reduce robot controller size. Journal of Multimodal User Interfaces. 2009;2(3-4):205–216.

Russell S, Norvig P. Artificial intelligence: a modern approach. 3rd ed. Upper Saddle River (NJ): Prentice Hall Press; 2009.

Salas E, Shuffler ML, Thayer AL, Bedwell WL, Lazzara EH. Understanding and improving teamwork in organizations: a scientifically based practical guide. In: Salas E, Rico R, Passmore J, editors. The Wiley Blackwell handbook of the psychology of team working and collaborative processes. New York (NY): Wiley, 2015.

Schaefer KE, Hill SG, Jentsch FG. Trust in human-autonomy teaming: a review of trust research from the US Army Research Laboratory. Robotics Collaborative Technology Alliance. In: Chen J, editor. AHFE 2018, AISC 784, Springer, 102–114; 2019. https://doi.org/10.1007/978-3-319-94346-6_10.

Schaefer KE, Chen JYC, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. Hum Factors. 2015;58(3):377–400.

Schaefer KE, Straub ER, Chen JYC, Putney J, Evans AW. Communicating intent to develop shared situation awareness and engender trust in human–agent teams. Sci Dir Cog Sys Res. 2017;46:26–39.

Scharre P. Army of none. New York (NY): Norton; 2018.

Selkowitz A, Lakhmani S, Larios C, Chen JYC. Agent transparency and the autonomous squad member. Proceedings of the 2016 International Annual Meeting of the Human Factors and Ergonomics Society; 2016 Sep 19–23; Washington DC.

Stowers K, Kasdaglis N, Rupp M, Chen J, Barber D, Barnes M. Insights into human–agent teaming: intelligent agent transparency and uncertainty. Proceedings of the Applied Human Factors and Ergonomics (AHFE) Conference; 2016 July 27–31; Orlando, FL.

Teo G, Einerman-Jones L, Barber D, Hudson I. Determining language for human to robot navigational commands. Proceedings of the Human Factors and Ergonomics Society meeting; 2014; Chicago, Illinois.

Vause N, Abouchacra K, Letowski T, Resta, E. Spatial (3-D) audio displays improve speech communication in the command and control vehicle (C2V). Proceedings of the 22nd Army Science Conference; 2000; Orlando, Florida. p. 834–841.

Vause N, Letowski T, Ferguson L, Mermagen T. Auditory issues in handheld land mine detectors. Proceedings of the Detection and Remediation Technologies for Mines and Minelike Targets IV; 1999 Aug 2; Orlando, FL. Bellingham, WA: SPIE 3710; c1999. p. 778–786.

Wang N, Pynadath D, Rovira E, Barnes MJ, Hill SJ. Is it my looks? Or something I said? The impact of explanations, embodiment, and expectations on trust and performance in human–robot teams. International Conference on Persuasive Technology; 2018: Springer; c2018. p. 56–69.

Wang N, Pynadath D, Hill S. Trust calibration within a human–robot team: comparing automatically generated explanations. Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction; 2016 Mar 7–10; New Zealand. p. 109–116.

Washington Times. Military services compete to increase relevance, funding amid Pentagon's shifting strategic focus. 2019 Mar 13 [accessed 2019 Aug 26]. https://www.washingtontimes.com/news/2019/mar/13/military-services-seek-more-relevance-funding-amid.

White T, Krausman A. Haas E. Tactile displays in Army operational environments. In: Martin J, Allender L, Savage-Knepshield P, Lockett J, editors. Designing soldier systems; current issues in human factors. Farnham, Surrey (England): Ashgate; 2012. p. 97–116.

Wickens CD. Multiple resources and mental workload. Hum Factors. 2008a;50(3):449–454.

Wickens CD. Situation awareness and workload in aviation. Curr DirPsychol Sci. 2008b;11(4):128–133.

Wickens CD. Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science. 2002;3(2):159–177.

Williams A, Scharre P, editors. Autonomous systems: issues for defence policymakers. DTIC TR. Norfolk (VA): Capability Engineering and Innovation Division Headquarters Supreme Allied Commander Transformation; 2015 Sep 30 [accessed 2019 Aug 29]. https://apps.dtic.mil/dtic/tr/fulltext/u2/1010077.pdf

Wright JL, Chen JY, Barnes MJ. Human–automation interaction for multiple robot control: the effect of varying automation assistance and individual differences on operator performance. Ergonomics. 2018;61(8):1033–1045.

Wright JL, Chen JY, Barnes MJ, Boyce MW. The effects of information level on human–agent interaction for route planning. Proceedings of the Human Factors and Ergonomics Society Annual Meeting; 2015;59(1):811–815.

Wright JL, Chen JY, Barnes MJ, Hancock PA. The effect of agent reasoning transparency on automation bias: an analysis of response performance. In: International Conference on Virtual, Augmented and Mixed Reality. Springer, Cham; 2016. p. 465–477.

Wright JL, Chen JY, Barnes MJ, Hancock PA. Agent reasoning transparency: the influence of information level on automation-induced complacency. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2017. Report No.: ARL-TR-8044.

Wright J, Chen JYC, Lakhmani S. Agent transparency and reliability in human-robot interaction: The influence on user confidence and perceived reliability. IEEE Transactions on Human–Machine Systems. Forthcoming 2019.

List of Symbols, Abbreviations, and Acronyms

2-D	two-dimensional
3-D	three-dimensional
AFRL	Air Force Research Laboratory
AI	artificial intelligence
AiTR	Aided Target Recognition
ANS	Autonomous Navigation Systems
ARL	Army Research Laboratory
ARPI	Autonomous Research Pilot Initiative
ASM	Autonomous Squad Member
CAST	Collaborative Agents for Simulating Teamwork
CCDC	Combat Capabilities Development Command
DOD	Department of Defense
DOF	depth of field
FOV	field of view
H-A	human participants
HAT	human–agent teaming
H-H	human–human
HRED	Human Research and Engineering Directorate
HRI	human–robot interactions
HSTAMIDS	Hand-held Stand-off Mine Detection System
IA	intelligent agent
ICT	Institute for Creative Technologies
IED	improvised explosive devices
IMPACT	Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies
MDS	mine detection system

ML	machine learning
NGCV (RV)	Next Generation Combat Vehicle robotic version
NLP	natural language processing
RCAST	RPD-enabled CAST
RPD	Recognition-Primed Decision-Making
SA	situation awareness
SAT	Situation awareness-based Agent Transparency
SID	smart interaction device
SMM	shared mental models
TNO	Toegepast Natuurwetenschappelijk Onderzoek
TOM	theory of mind
UAS	Unmanned Aerial Systems
UCF	University of Central Florida
UGV	Unmanned ground vehicles
XAI	eXplainable AI

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

1 (PDF) CCDC ARL
FCDD RLD CL
TECH LIB

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) CCDC ARL
FCDD RLH B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

1 (PDF) USA ARMY G1
DAPE HSI M SAMS
300 ARMY PENTAGON
RM 2C489
WASHINGTON DC 20310-0300

1 (PDF) USAF 711 HPW
711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH
45433-7604

1 (PDF) USN ONR
ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 (PDF) USA NSRDEC
RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 (PDF) OSD OUSD ATL
HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

<u>ABERDEEN PROVING GROUND</u>

11 (PDF) CCDC ARL
FCDD RLH
J LANE
Y CHEN
P FRANASZCZUK
K MCDOWELL
K OIE
FCDD RLH BD
D HEADLEY
M BARNES
FCDD RLH FA
A DECOSTANZA
FCDD RLH FB
A EVANS
FCDD RLH FC
J GASTON
FCDD RLH FD
A MARATHE